# IV. Capture-Recapture Models with Individual Covariates

# INDIVIDUAL COVARIATE MODELS

- p depends on an explicit covariate which you only observe the values of for individuals that are encountered

- logit(p[i]) = alpha0 + alpha1*x[i]

- For this reason, called "Model Mx" (Kery and Schaub BPA)

- Two diametrically opposite approaches to analysis
  - (1) Horvitz-Thompson estimation (Huggins and Alho used this idea based on conditional likelihood)

  - (2) Model-based "full likelihood": Put a distribution on x[i] (Borchers et al. 1998; Royle 2009, Biometrics)

# Model Mx

**We still have a binomial encounter model:**

y[i] ~ Binomial(p[i], K)

logit(p[i]) = a + b*x[i]

- x[i] is NOT OBSERVED for uncaptured individuals

**Model for the covariate:**

x[i] ~ normal($\mu_x$, $\sigma_x$)  (or similar)

# INDIVIDUAL COVARIATES

- Huggins, R. M. (1989). On the statistical analysis of capture experiments. *Biometrika*, 76(1), 133-140.

- Alho, J. M. (1990). Logistic regression in capture-recapture models. *Biometrics*, 46(3), 623.

- Borchers, D. L., Zucchini, W., & Fewster, R. M. (1998). Mark-recapture models for line transect surveys. *Biometrics*, 1207-1220.

- Pollock, K. H. (2002). The use of auxiliary variables in capture-recapture modelling: an overview. *Journal of Applied Statistics*, 29(1-4), 85-102.

- Royle, J. A. (2009). Analysis of capture–recapture models with individual covariates using data augmentation. *Biometrics*, 65(1), 267-274.

- Borchers, D.L., S.T. Buckland and W. Zucchini. 2002. Estimating Animal Abundance: Closed Populations (Springer Verlag) (Ch. 11).

# Model Mx and multi-session models

- **Conceptually and technically Model Mx is exactly like "class-structured" models considered previously except Model Mx usually used in context of a continuous covariate.**

- Consider having sex-specificity of model parameters

logit(p[i]) = alpha0 + alpha1*Xsex[i]

  Xsex[i] =  0 if female
               1 if male

- Xsex is missing for M-n individuals in our augmented data set.

- Put a prior distribution on it….With a discrete covariate, the prior is "class membership"

# Example of model Mh

- Microtus data from Williams et al. (2002)

```
source("microtus.data.R")
 head(microtus.data)
     [,1] [,2] [,3] [,4] [,5] [,6]
[1,]    0    1    1    0    0   37
[2,]    1    0    1    1    0   46
[3,]    1    1    1    0    1   60
[4,]    1    1    1    1    1   49
[5,]    0    0    0    0    1   38
[6,]    1    0    1    0    1   40
```

$n = 56$    $K = 5$ sample occasions, $x[i]$ = "body mass" is stored in column 6 of the matrix

Model for $x[i]$: $x[i] \sim$ Normal(mu, sigma)

# Analysis of the Microtus data

- R work session

R script: `closed_models_part4.R`

# Model Mx: toward SCR

Model Mx has been widely adopted for estimation of N in capture-recapture studies to account for spatial heterogeneity in encounter probability, by defining:

    x = "distance to edge" (DTE)

This is estimated for each captured individual and treated as fixed and known.

Boulanger and McLellen (2001)
Ivan and White (2013)

Problems: variable precision. Biased near edge! Also doesn't account for trap-level information.

# Applying model Mx

- We'll use x = distance to centroid (DTC) of trap array, call this "xdist":

    xdist[i] = dist(sbar[i], x0)

    x0 = mean trap location (centroid of trap array)
    sbar[i] = average location of individual i.

To do the Bayesian analysis by DA we need a prior for xdist[i] to account for uncaptured individuals. Could as well just put the prior on sbar[i] since x0 is known.

# Prior for d[i] or prior for s[i]?

- **Prior for xdist[i]:** What are the possible values for where captureable individuals come from?

- xdist[i] ~ dunif(0, Dmax)

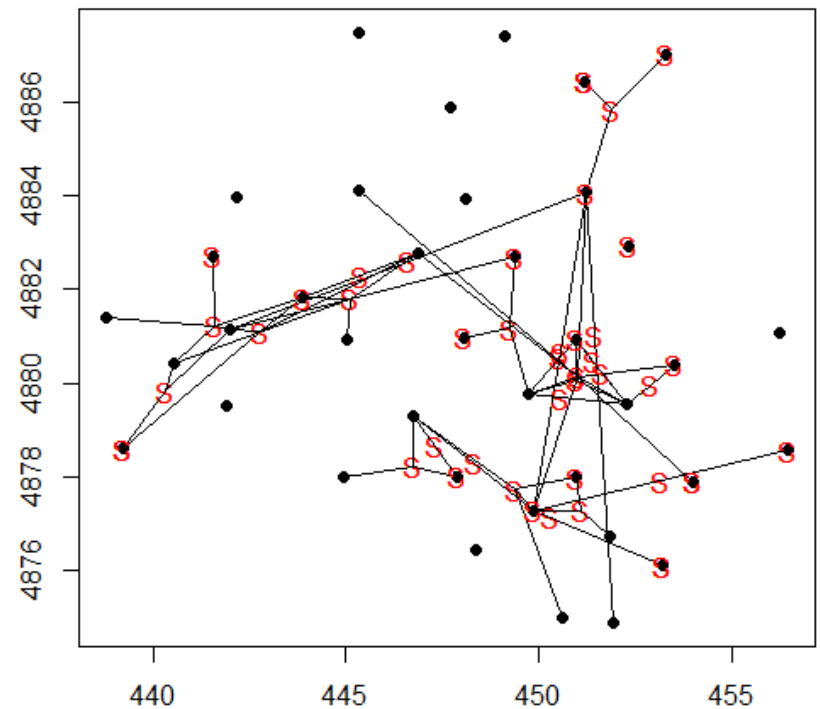    Dmax = furthest possible capture?

# Fort Drum black bear data

```
library("scrbook")
data(beardata)
nind<-dim(beardata$bearArray)[1]
K<-dim(beardata$bearArray)[3]
ntraps<-dim(beardata$bearArray)[2
toad<- spiderplot(beardata$bearAr

## Distance to centroid of traps
xdist<-toad$xcent

## average location of capture
sbar<- toad$avg.s

## Centroid of trap array
x0<- toad$center
```

# ESTIMATING DENSITY

- **By putting a prior distribution on xdist[i] this explicitly defines an AREA within which the sampleable population lives. That is, N is all individuals within Dmax of the centroid**
- You will find that the estimated N changes as you change Dmax.

- The DTC/DTE model, with a prior on the distance covariate, simultaneously estimates N and Density

- Provides resolution to the "unknown area" problem

- (was **not** noted by Boulanger and McLellan 2001)

# MODEL MX HAS PROBLEMS

(1)  Subjective choice of Dmax has a big effect – this model implies that density of individuals decreases as you move away from the centroid

Area of concentric rings INCREASES as you move away. So a constant frequency of individuals corresponds to lower density

(2) Use of estimated DTC (or DTE) is biased and estimated with variable precision. Model does not account for that.

# IMPROVING MODEL MX

- **Improvement 1**: Instead of messing with a prior on xdist[i] why not just put the prior on sbar[i]? Exactly the same model, just a different prior (via a transformation).

- **Improvement 2**: Instead of distance to x0, why not distance to each trap x[j]?

  xdist[i,j] = dist(sbar[i], x[j])

- **Improvement 3**: "sbar" is really a surrogate for "center of activity" – which is unobserved. Make it a latent variable (like Model Mh but with some indirect information)

# Improvement 1 of model Mx

Improvement 1: Instead of messing with a prior on xdist[i] why not just put the prior on sbar[i]? **Exactly the same model, just a different prior (via a reparameterization**).

- sbar[i] is the average capture location. But it's really a surrogate for "where individual i lives" – home range center?

- What kind of prior makes sense for this? In the absence of information, how about sbar[i] ~ Uniform(space around trap array) ???
  
  S = "space around the trap array"

# The prior distribution for "sbar" for the Fort Drum model
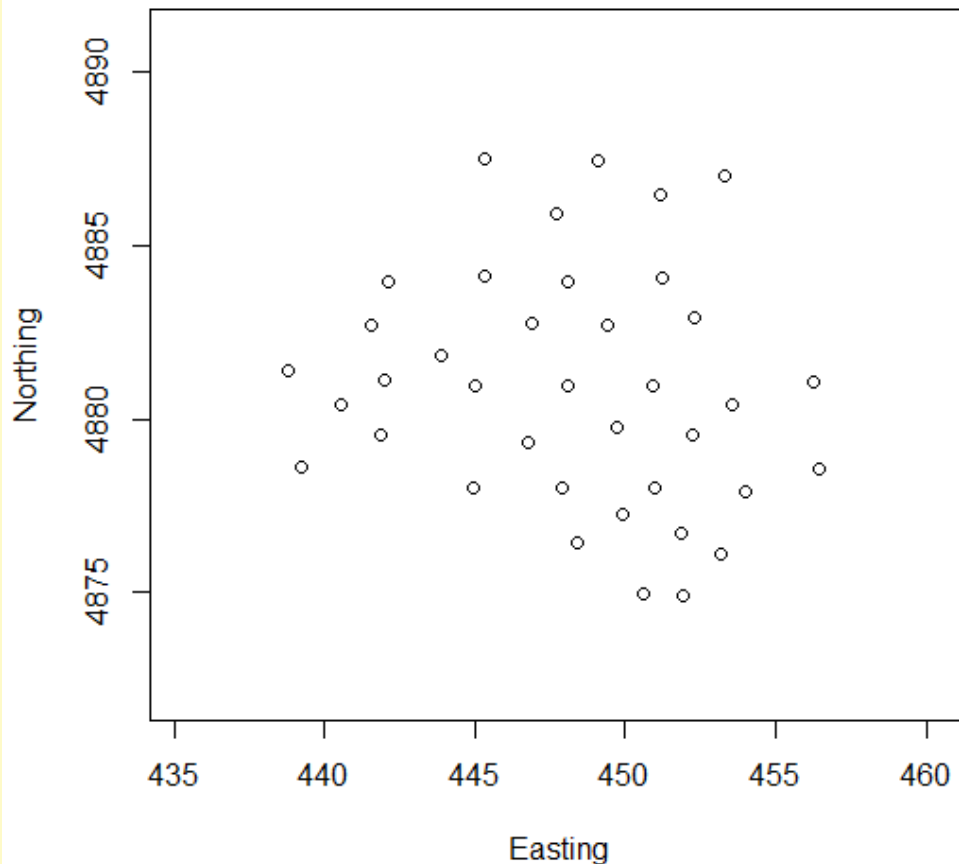


- sbar is Uniform(S)

- S defined by: 4 unit buffer around the minimum and maximum x- and y-coordinates

- Try different buffers and verify stabilizing Density = N/area

# A WARNING ABOUT COORDINATE SCALING IN BUGS

- The model which regards sbar as a variable effectively predicts sbar for each **uncaptured individual**

- WinBUGS seems to only carry around 4 significant digits (or else R2WinBUGS rounds to 4 digits, we're not sure).

- The coordinates of sbar for Fort Drum is 4xx.x and 48x.x – the leading 4 and 48 are costing precision for estimating the coordinates of uncaptured bears due to this 4-digit truncation

- JAGS does not appear to suffer this problem

- If we use BUGS it is imperative that we scale/translate the coordinate system so that we're not carrying around unnecessary digits (or use JAGS)

# Improvement 2 of model Mx

Improvement 2: Instead of distance to centroid, why not distance to each trap x[j]?

$xdist[i,j] = dist(sbar[i], x[j])$

$logit(p[i,j]) = alpha0 + alpha1*xdist[i,j]$

Note: p now depends on i and j

Traps are just replicate sample occasions, like distinct methods, or sample frames, or observers, but with trap-specific p.

# Improvement 3 of model Mx

- sbar is meant to be an estimate of something, say s, the centroid of activity – "activity center", home range center, etc..

- s is strictly unknown. Regard it as a latent variable.

- In BUGS: input initial values for it, remove from "data"

- No plug-in estimation bias and heterogeneous variance.

# SUMMARY OF CAPTURE-RECAPTURE PART 4

- Individual covariate models with distance-to-edge/distance-to-centroid resolve some technical problems with ordinary CR models
  - Heterogeneity in p related to variable exposure to trapping
  - Putting a distribution on the covariate resolve "unknown area"

- Useful as a starting point for developing fully spatial capture-recapture models
  - Model location instead of distance
  - Distance to each trap
  - Treat "s" as a latent variable